

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-301953

(43)Date of publication of application : 13.11.1998

(51)Int.Cl.

G06F 17/30

G06T 1/00

(21)Application number : 09-111504

(71)Applicant : JUST SYST CORP

(22)Date of filing : 28.04.1997

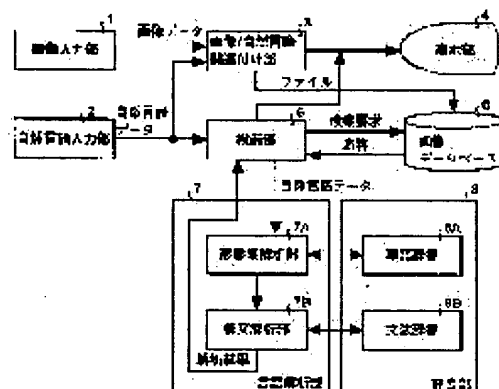
(72)Inventor : UKIKAWA KAZUNOBU  
UKIKAWA HATSUKO

(54) IMAGE MANAGING DEVICE, IMAGE RETRIEVING DEVICE, IMAGE MANAGING METHOD, IMAGE RETRIEVING METHOD, AND COMPUTER-READABLE RECORDING MEDIUM RECORDING PROGRAM FOR ALLOWING COMPUTER TO EXECUTE THESE METHODS

(57)Abstract:

PROBLEM TO BE SOLVED: To more surely and efficiently retrieve a required image.

SOLUTION: Image data and natural language data indicating the visual impression of an image formed based on the image data are relationally stored in an image data base 5. In the case of executing retrieval by a retrieving part 6, natural language data indicating a retrieving sentence are inputted from a natural language input part 2 and collated with the natural language data stored in the data base 5, so that image data related to the natural language data having a common visual impression are extracted from the data base 5.



## LEGAL STATUS

[Date of request for examination] 03.06.1998

[Date of sending the examiner's decision of rejection] 21.12.1999

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

BEST AVAILABLE COPY

## \* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]

[Claim 1] Image management equipment characterized by associating and memorizing the natural language data in which the visual impressions about the image formed based on image data and this image data are shown in the image management equipment which puts in a database and manages image data.

[Claim 2] It has the image database which associates and memorizes the natural language data in which the visual impressions about the image formed based on image data and this image data are shown in the image management equipment which puts in a database and manages image data. Said image database is accessed using a certain visual impressions. Image management equipment characterized by answering the image data relevant to natural language data including said a certain visual impressions when said a certain visual impressions are some visual impressions shown by the natural language data on a self-database.

[Claim 3] Furthermore, image management equipment according to claim 1 or 2 characterized by relating with image data the natural language data which have an input means to input natural language data through a key stroke, speech recognition, image recognition, etc., and were inputted by this input means, and memorizing them.

[Claim 4] Said image data is image management equipment of any one publication of claim 1-3 characterized by being data, such as still picture data and a video data.

[Claim 5] The image database which associates and memorizes the natural language data in which the visual impressions about the image formed based on image data and this image data are shown, The natural language data inputted by said retrieval sentence input means to the natural language data remembered to be a retrieval sentence input means to input the natural language data in which a retrieval sentence is shown by said image database are collated. Image retrieval equipment characterized by having a retrieval means to extract the image data relevant to the natural language data with which visual impressions are common according to the collating result from said image database.

[Claim 6] Said retrieval means is image retrieval equipment according to claim 5 characterized by collating after carrying out language analysis of the natural language data inputted by the natural language data memorized by said image database and said retrieval sentence input means.

[Claim 7] Said retrieval means is image retrieval equipment according to claim 5 or 6 characterized by extracting two or more image data from the higher one of similarity in ascending order according to said collating result.

[Claim 8] The information database which supports inference for drawing new visual impressions from two or more visual impressions, The image database which associates and memorizes the natural language data in which two or more visual impressions about the image formed based on image data and this image data are shown, About each natural language data remembered to be a retrieval sentence input means to input the natural language data in which a retrieval sentence is shown by said image database, referring to said information database The inference means which draws the new visual impressions equivalent to the visual impressions of the natural language data inputted by said retrieval sentence input means from two or more visual impressions, A retrieval means to extract the image data relevant to these natural language data from said image database when there are natural language data which were able to draw said new visual impressions with said inference means among the natural language data memorized by said storage means, Image retrieval equipment characterized by preparation \*\*\*\*\*.

[Claim 9] Said inference means is image retrieval equipment according to claim 8 characterized by what is reasoned after carrying out language analysis of the natural language data inputted by the natural language data memorized by said image database and said retrieval sentence input means.

[Claim 10] Furthermore, image retrieval equipment of any one publication of claim 5-9 characterized by having an image display means to perform image display based on the image data extracted by said retrieval means.

[Claim 11] The image management method characterized by including the 1st process which associates the natural language data in which the visual impressions about the image formed based on image data and this image data are

shown in the image management method which puts in a database and manages image data, and the 2nd process which memorizes the image data and natural language data which were associated by said 1st process.

[Claim 12] The 1st process which inputs the natural language data in which a retrieval sentence is shown, The 2nd process which collates the natural language data inputted according to said 1st process to the natural language data memorized by the image database which associated and memorized the natural language data in which the visual impressions about the image formed based on image data and this image data are shown, The image search method characterized by including the 3rd process which extracts the image data relevant to the natural language data with which visual impressions are common according to the collating result of said 2nd process from said image database.

[Claim 13] The record medium which is characterized by recording the program which makes a computer perform the approach indicated by said claims 11 or 12 and in which computer reading is possible.

---

[Translation done.]

**\* NOTICES \***

**JPO and NCIPi are not responsible for any damages caused by the use of this translation.**

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the record medium which recorded the program which makes a computer perform the image retrieval equipment with which an image is searched using the image management equipment which manages an image using natural language and its approach, and natural language, its approach, and those approaches and in which computer reading is possible.

[0002]

[Description of the Prior Art] Conventionally, the technique of searching an image using natural language is proposed with JP,4-180175,A, the 4-264972 official report, the 6-176121 official report, the 7-334507 official report, the 8-138075 official report, etc. According to each official report, the content is interpreted by performing language analyses, such as morphological analysis, about the inputted natural language, and the technique of searching the image suitable for the content is indicated.

[0003] In this technique, attribute information, such as a profile of that image, relates with an image beforehand, it memorizes, and a desired image can be acquired by testing the result of language analysis by comparison to attribute information.

[0004] Moreover, as attribute information which relates with an image not only in the technique by the official report mentioned above, but generally, and is memorized, there are time of filing, an operator's name, classification (a photograph, a picture, document, etc.) of an image, etc., and, as for a user, the hand of retrieval uses these attributes information as a loan.

[0005]

[Problem(s) to be Solved by the Invention] Like the above-mentioned official report, since the attribute information which serves as a search key to an image was added beforehand and the attribute information was the information of almost objective and necessary minimum although the desired image could be obtained when giving the natural language relevant to the attribute information on the occasion of retrieval, the technique by the conventional example had a possibility that a user might forget as long as there is nothing by the memorandum. So, since he has forgotten the attribute information from which a user becomes a search key, or possibility of giving the natural language data of the content which separated from the attribute information which a desired image has become large on the occasion of retrieval when skeptical about the attribute information in storage, finding out a desired image out of a vast quantity of images had the problem of being difficult.

[0006] This invention aims at obtaining the record medium which recorded the program which makes a computer perform the image management equipment which is easy to access a desired image regardless of attribute information, its approach, and its approach and in which computer reading is possible in order to solve the problem by the conventional example mentioned above.

[0007] Moreover, this invention sets it as other objects to obtain the record medium which recorded the program which makes a computer perform the image retrieval equipment which can search a desired image more certainly and efficiently, its approach, and its approach and in which computer reading is possible in order to solve the problem by the conventional example mentioned above.

[0008] In addition, as an approximation technique, like JP,6-131403,A, although there are some by which the dictionary which associated natural language, image information, and speech information as an item, and registered them to the word (keyword) is indicated, the correlation with natural language and an image is not technically indicated about retrieval of an image.

[0009]

[Means for Solving the Problem] In order to solve the technical problem mentioned above and to attain the object, the image management equipment concerning invention of claim 1 is characterized by associating and memorizing the natural language data in which the visual impressions about the image formed based on image data and this image data are shown in the image management equipment which puts in a database and manages image data.

[0010] Since the natural language data in which the visual impressions of the image formed based on image data and its image data are shown were associated and memorized according to this claim 1, it is possible to put in a database with the subjective element in which visual impressions have an image.

[0011] Moreover, the image management equipment concerning invention of claim 2 It has the image database which associates and memorizes the natural language data in which the visual impressions about the image formed based on image data and its image data are shown in the image management equipment which puts in a database and manages image data. An image database is characterized by answering the image data relevant to natural language data including a certain visual impressions, when it is some visual impressions where it is accessed using a certain visual impressions, and a certain visual impressions are shown by the natural language data on a self-database.

[0012] According to this claim 2, the natural language data in which the visual impressions of the image formed in an image database based on image data and its image data are shown are associated and memorized. Since the image data relevant to the natural language data was made to answer when access was performed by the natural language data of a certain visual impressions on an image database Regardless of attribute information, access by visual impressions, i.e., a subjective impression, is attained, and it becomes easy to access a desired image by this.

[0013] Moreover, in invention of claims 1 or 2, the image management equipment concerning invention of claim 3 has further an input means to input natural language data through a key stroke, speech recognition, image recognition, etc., and is characterized by relating with image data the natural language data inputted by the input means, and memorizing them.

[0014] It is possible to give the visual information which has subjectivity in each image, without being caught by the objectivity of an image, since the natural language data for relating with image data and memorizing were inputted through a key stroke, speech recognition, image recognition, etc. according to this claim 3.

[0015] Moreover, the image management equipment concerning invention of claim 4 is characterized by image data being data, such as still picture data and a video data, in any one invention of claims 1-3.

[0016] According to this claim 4, since image data was used as data, such as still picture data and a video data, it is possible to give the visual information on arbitration to each image regardless of the classification of an image.

[0017] Moreover, the image retrieval equipment concerning invention of claim 5 The image database which associates and memorizes the natural language data in which the visual impressions about the image formed based on image data and image data are shown, The natural language data inputted by the retrieval sentence input means to the natural language data remembered to be a retrieval sentence input means to input the natural language data in which a retrieval sentence is shown by the image database are collated. It is characterized by having a retrieval means to extract the image data relevant to the natural language data with which visual impressions are common according to the collating result from an image database.

[0018] According to this claim 5, the natural language data in which the visual impressions about the image formed in an image database based on image data and its image data are shown are associated and memorized. By inputting the natural language data in which a retrieval sentence is shown, and collating with the natural language data on an image database Since the image data relevant to the natural language data with which visual impressions are common was extracted from the image database It is possible for an image to be searched with visual impressions, i.e., a subjective impression, even if it does not know the attribute information on an image, and to search a desired image more certainly and efficiently by this.

[0019] Moreover, the image retrieval equipment concerning invention of claim 6 is characterized by collating it, after a retrieval means carries out language analysis of the natural language data inputted by the natural language data and the retrieval sentence input means which were memorized by the image database in invention of claim 5.

[0020] Since it was made to collate according to this claim 6 after carrying out language analysis of the natural language data inputted as the natural language data memorized by the image database on the occasion of retrieval, high collating of the precision suitable for the semantics of visual impressions is attained.

[0021] Moreover, as for a retrieval means, the image retrieval equipment concerning invention of claim 7 is characterized by extracting two or more image data from the higher one of similarity in ascending order according to a collating result in invention of claims 5 or 6.

[0022] According to this claim 7, since ascending order extracted two or more image data from the higher one of similarity according to the collating result on the occasion of retrieval, it is possible for two or more images near visual

impressions to be mentioned as a candidate, and to search a desired image more efficiently even from visual impressions by this.

[0023] Moreover, the image retrieval equipment concerning invention of claim 8 The information database which supports inference for drawing new visual impressions from two or more visual impressions, The image database which associates and memorizes the natural language data in which two or more visual impressions about the image formed based on image data and this image data are shown, About each natural language data remembered to be a retrieval sentence input means to input the natural language data in which a retrieval sentence is shown by the image database, referring to an information database The inference means which draws the new visual impressions equivalent to the visual impressions of the natural language data inputted by the retrieval sentence input means from two or more visual impressions, When there are natural language data which were able to draw new visual impressions with the inference means among the natural language data memorized by the storage means, it is characterized by having a retrieval means to extract the image data relevant to these natural language data from an image database.

[0024] Associating and memorizing the natural language data in which two or more visual impressions about the image formed in an image database based on image data and its image data are shown, and referring to an information database according to this claim 8 When the new visual impressions equivalent to the visual impressions of the natural language data which are a retrieval sentence are able to be drawn from two or more visual impressions of a certain natural language data on an image database Since the image data relevant to the natural language data of a certain was extracted from the image database It is possible to be able to secure as an object for an extract of an image certainly, even if it is natural language data which are overlooked, when the visual information on an image database is used independently, and to search a desired image more certainly and efficiently by this in the case of retrieval.

[0025] Moreover, the image retrieval equipment concerning invention of claim 9 is characterized by reasoning it, after an inference means carries out language analysis of the natural language data inputted by the natural language data and the retrieval sentence input means which were memorized by the image database in invention of claim 8.

[0026] Since it was made to reason according to this claim 9 after carrying out language analysis of the natural language data inputted as the natural language data memorized by the image database on the occasion of inference, high inference of the precision suitable for the semantics of visual impressions is attained.

[0027] Moreover, the image retrieval equipment concerning invention of claim 10 is characterized by having an image display means to perform image display further based on the image data extracted by the retrieval means in any one invention of claims 5-9.

[0028] Since it was made to perform image display based on the image data extracted by retrieval according to this claim 10, it is possible for a retrieval result to be shown visually and to check the correction of a retrieval result easily by this.

[0029] Moreover, the image management method concerning invention of claim 11 is characterized by including the 1st process which associates the natural language data in which the visual impressions about the image formed based on image data and this image data are shown, and the 2nd process which memorizes the image data and natural language data which were associated by the 1st process in the image management method which puts in a database and manages image data.

[0030] Since according to this claim 11 it was made the process which memorizes image data and natural language data after associating the natural language data in which the visual impressions of the image formed based on image data and this image data are shown, it is possible to put in a database with the subjective element in which visual impressions have an image.

[0031] Moreover, the image search method concerning invention of claim 12 The 1st process which inputs the natural language data in which a retrieval sentence is shown, The 2nd process which collates the natural language data inputted according to the 1st process to the natural language data memorized by the image database which associated and memorized the natural language data in which the visual impressions about the image formed based on image data and this image data are shown, It is characterized by including the 3rd process which extracts the image data relevant to the natural language data with which visual impressions are common according to the collating result of the 2nd process from an image database.

[0032] According to this claim 12, the natural language data in which a retrieval sentence is shown are inputted. The natural language data inputted to the natural language data memorized by the image database which associated and memorized the natural language data in which the visual impressions about the image formed based on image data and its image data are shown are collated. Since the image data relevant to the natural language data with which visual impressions are common according to the collating result was made into the process extracted from an image database It is possible for an image to be searched with visual impressions, i.e., a subjective impression, even if it does not know the attribute information on an image, and to search a desired image more certainly and efficiently by this.

[0033] Moreover, the record medium concerning invention of claim 13 is having recorded the program which makes a computer perform the approach indicated by claims 11 or 12, and reading becomes possible by computer about the program, and it can realize actuation of claims 11 or 12 by computer this.

[0034]

[Embodiment of the Invention] The gestalt of suitable operation of the record medium which recorded the program which makes a computer perform the image management equipment applied to this invention with reference to an accompanying drawing below, image retrieval equipment, an image management method, image search methods, and those approaches and in which computer reading is possible is explained to a detail.

[0035] (Gestalt 1 of operation) A principle is explained first. Drawing 1 - drawing 3 are drawings explaining the principle by the gestalt 1 of implementation of this invention. The image with which the glass into which the liquid to which the color (it considers as red) was attached was put is apparently placed on the circle table of 1 piece is shown in drawing 1 . Supposing human being holds a subjective impression to this image, it is not an exception to hold the following impressions, either. That is, "it seems to be high if the liquid in a glass is wine" (it considers as the visual information 1), "it seems that the liquid to which the color was attached is in a glass as for close. although it is not exact -- wine -- like . " (it considers as the visual information 2) -- "-- it seems that this drink has got cold although it is not exact " (it considers as the visual information 3) -- the glass containing "liquid is placed on the table. " (it considers as the visual information 4) -- \*\* -- it is an impression [ like ].

[0036] Although the above-mentioned visual information 1 and 2 does not have the information which specifies wine specially in the image of drawing 1 , it expresses the impression acquired from the point that the color of a liquid is red, and the point that the configuration of a glass resembles the configuration of a wineglass. The above-mentioned visual information 3 expresses the impression on which the color of the liquid which is in a glass appeared by imagining the drink which could not imagine a drink hot from the drink which carried out red since it was red, but rather got cold like wine or virgin bloody Mary. The above-mentioned visual information 4 expresses the impression about the physical relationship of a table and a glass simply.

[0037] When four kinds of above visual information 1-4 is compared, although the visual information 1-3 is subjective since the visual information 4 is rather expressing the composition of an image as it is, although it corresponds to a subjective impression since the visual impression (semantics is also included also as a guess) received from an image is expressed as a description of the image, it is close to an objective impression.

[0038] the physical relationship and the table of components in the image that the glass be place on a table by the visual information 1-3 be a circle table -- etc. -- although the impression about a configuration be include , if it be going to mention the description without see the image of drawing 1 later , since people have remember [ more ] by the fragmentary impression generally , they will be expect that an expression like the visual information 1-3 be use .

[0039] This invention defines as a file what associated visual information and image data using a subjective impression like the visual information 1-4 remaining to later to an image. That is, if the image of drawing 1 is mentioned as an example, the file of an image will serve as a configuration which related the visual information 1-4 with the image in the form of natural language, as shown in drawing 2 . Therefore, when searching the image of drawing 1 from two or more images, a desired image ( drawing 1 ) can be specified from two or more images by the retrieval sentence that what is necessary is just to give a retrieval sentence with the natural language of the expression near the visual information 1-4 mentioned above.

[0040] Then, as a premise of retrieval, as shown in drawing 3 , Image a (image of drawing 1 ), Image b, and Image c shall exist. The above-mentioned visual information 1-4 is related with Image a, the visual information on "it seems "it seems to be Biel which got cold well", that a toast is given in the beer mug although it is not exact, etc." is related with Image b, and visual information, like "it seems "it seems that coffee has placed on a table", that steam has come out from a cup although it is not exact, etc." is related with Image c.

[0041] In order to search the desired image a from two or more above images a, b, and c, when a retrieval sentence "the image which has placed the drink which got cold on the table is ?" is given, as shown in drawing 3 R> 3, the visual information (natural language) related with the retrieval sentence and each images a, b, and c is collated. An image with the highest visual information on similarity is extracted as an object of retrieval by this collating.

[0042] A retrieval sentence becomes the point of collating of the impression which shows the physical relationship of a table and a drink, and the impression that the drink has got cold. In the case of Image a, the impression "the drink has got cold" from the visual information 3 corresponds to the point of a retrieval sentence, and the impression of "having placed the glass containing a liquid on the table" from the visual information 4 corresponds to it.

[0043] That is, Image a matches the image which has placed the drink which got cold on the table. Moreover, in the case of Image b, only the impression of the drink (Biel) which got cold well to the point of a retrieval sentence corresponds,



and in being Image c, only the impression of the physical relationship of having placed the cup on the table to the point of a retrieval sentence corresponds.

[0044] Therefore, in the example of drawing 3, the image a with the highest visual information on similarity is extracted as an object of retrieval, and the other images b and c become the outside of an object.

[0045] Below, it explains using concrete equipment. Drawing 4 is the block diagram showing functionally the image processing system by the gestalt 1 of implementation of this invention. The image processing system shown in drawing 4 has the function of both the image management equipment and image retrieval equipment concerning this invention.

[0046] The image processing system shown in drawing 4 is equipped with the image input section 1, the natural language input section 2, the image / natural language correlation section 3, a display 4, an image database 5, the retrieval section 6, the language analysis section 6, and the dictionary section 7. The image input section 1 inputs the image data which should be registered into an image database 5, and the natural language input section 2 inputs the visual information which is associated and is added to the image data which should be registered into an image database 5, and the retrieval sentence at the time of retrieval as natural language data.

[0047] An image / natural language correlation section 3 associates the natural language data inputted by the image data inputted by the image input section 1 and the natural language input section 2, and outputs them to an image database 5 by considering the associated data as a file. Since this image / natural language correlation section 3 relate displaying an image and visual information, it outputs the indicative data based on image data and natural language data to a display 4.

[0048] In the case of retrieval, a display 4 receives an indicative data from the retrieval section 6, and displays a retrieval sentence and a retrieval result (image) while it receives an indicative data from an image / natural language correlation section 3 and displays an image and visual information in the case of correlation. An image database 5 answers natural language data and image data according to the retrieval demand of the retrieval section 6 while memorizing the file to which the image data associated by an image / natural language correlation section 3 and natural language data were made to correspond (reply).

[0049] Since language analysis of an input retrieval sentence and the visual information on an image database 5 is carried out by the language analysis section 7 according to the retrieval sentence (natural language data) inputted by the natural language input section 2, the retrieval section 6 is collated. This retrieval section 6 extracts the image data related with the high visual information on similarity to the input retrieval sentence according to that collating result from an image database 5, and displays that retrieval result (extracted image) on a display 4.

[0050] Although it is an example, well-known natural language processing is already applicable to the natural-language-processing section by the language analysis section 7 and the dictionary section 8. As reference, there are the Institute of Electronics, Information and Communication Engineers and 1988 "the basic technique of natural language processing" (Nomura \*\*\*\*\*).

[0051] The language analysis section 6 is constituted by morphological analysis section 7A and syntax analyzer 7B, and the dictionary section 8 is constituted by word dictionary 8A which supports morphological analysis, and syntax dictionary 8B which supports syntax analysis. Morphological analysis section 7A performs morphological analysis based on the natural language data inputted with reference to word dictionary 8A, and syntax analyzer 7B analyzes syntactic structure based on the natural language data inputted with reference to syntax dictionary 8B. The analysis result of this syntax analyzer 7B is outputted to the retrieval section 6.

[0052] Then, an image database 5 is explained. Drawing 5 is drawing which explains roughly the content of storage of the image database 5 by the gestalt 1 of operation. The image database 5 has memorized the image data and natural language data which were made to correspond to a file name and were associated, as shown in drawing 5. Generally the file name is mentioned here, because the file name of a proper is defined as one file, and the file name itself is not related to the meaning of this invention.

[0053] For example, as for the file of file name FILEA, image data IMGA and the natural language data NLDA are associated, and, as for the file of file name FILEB, image data IMGB and the natural language data NLDB are associated. From making visual information into the content, the natural language data contained in each file can give two or more visual information like the natural language data NLDA, for example.

[0054] Plurality here does not show the semantics in which it is shown in which that two or more sentences expressing a visual impression exist, and two or more visual information that it became independent exists. That is, although it will be classified into two or more visual information if the context is interpreted, as natural language data, he is one lump. Thus, giving visual impressions from all views indefinitely to one image is allowed.

[0055] Below, actuation is explained. First, the Maine actuation is explained. Drawing 6 is a flow chart explaining the Maine actuation of the gestalt 1 of operation, and drawing 7 is drawing showing an example of the display transition at the time of the Maine actuation of the gestalt 1 of operation.



[0056] As actuation, the menu (refer to drawing 7 (a)) which made alternative the termination icon 103 for performing the retrieval processing icon 102 for operating retrieval of the filing processing icon 101 for operating filing of an image and an image and termination actuation is first displayed on the screen of a display 4 (step S1). Then, it is sensed whether which icon of the filing processing icon 101, the retrieval processing icon 102, and the termination icon 103 is chosen. This selection actuation shall be performed by the non-illustrated control unit.

[0057] And as shown in drawing 7 (b), when the filing processing icon 101 is chosen (the slash in drawing shows), (step S2) and processing shift to filing processing of drawing 8 (step S5). Then, processing returns to the menu display of step S1 again. Moreover, as shown in drawing 7 (c), when the retrieval processing icon 102 is chosen (the slash in drawing shows), (step S3) and processing shift to retrieval processing of drawing 12 (step S6). Then, processing returns to the menu display of step S1 again. Moreover, when the termination icon 103 is chosen, (step S4) and this processing are ended.

[0058] Then, filing actuation of an image is explained. Drawing which explains roughly the example of storage of the image database by the gestalt 1 of operation, and drawing 11 are [ drawing 9 / the flow chart with which drawing 8 explains filing actuation of the gestalt 1 of operation, and ] drawing showing an example of the display transition in the animation at the time of filing actuation of the gestalt 1 of operation in drawing and drawing 10 which show an example of the display transition at the time of filing actuation of the gestalt 1 of operation.

[0059] If processing shifts to filing processing, image data will be first inputted by the image input section 1 (step S51), and an image will be displayed on a display 4 based on the input image data (step S52). As shown in drawing 9 (a), specifically, the image display area 201 and the natural language display area 202 are displayed on the screen of a display 4. Here, the above-mentioned image a (refer to drawing 3 ) is displayed on the image display area 201 as an example.

[0060] Furthermore, as shown in drawing 9 R> 9 (a), the message (at the example of drawing, it is "inputting visual information") 203 of a purport which inputs the visual information about the image a currently displayed on the image display area 201 with natural language is displayed on the natural language display area 202 (step S53). In addition, in order to carry out termination actuation of this filing processing, the termination icon 204 is displayed in the natural language display area 202.

[0061] Then, visual information, i.e., a natural language entry of data, is received (step S54). Natural language data are data inputted through the natural language input section 2. Initiation of a natural language input displays visual information like "the liquid in a glass was wine" in the natural language display area 202, as shown in drawing 9 (b). In that case, it is displayed as Cursor CUR shows the input location and correction location of the following, and migration actuation of this cursor CUR is carried out by the control unit which is not illustrated.

[0062] And if the input of visual information "it has placed on the liquid -- table in a glass." is completed as shown in drawing 9 (c), input termination will be checked by selection actuation of the termination icon 204 (step S55). Therefore, processing of step S54 and step S55 is repeatedly performed until selection actuation of this termination icon 204 is carried out.

[0063] Thus, after the input of visual information is completed (step S55), a response with the natural language data for forming the visual information currently displayed on the image data and the natural language display area 202 for forming the image a currently displayed on the image display area 201 is associated, and as shown in drawing 10 as one file, registration to an image database 5 is performed (step S56). actuation of the control unit which is not illustrated in this registration -- "\*\*\*\*\* -- a file name can be given like that 1." In addition, although it is possible to also add the attribute information on other, since it is not related to the meaning of this invention, explanation is omitted here.

[0064] In the case of drawing 9 , although a still picture (for example, photograph) like Image a is mentioned as the example, in the gestalt 1 of this operation, processing with the same said of continuous image, i.e., animation, is performed.

[0065] That is, visual information can be given, looking at change of a continuous frame, in using a series of gymnastic actuation as an animation, as shown in drawing 11 . For example, see the hit of a frame 1 and visual information like "turning a shoulder, after raising both hands upwards" is inputted. Next, both guide pegs are opened and -- waist is rotated. further -- the hit Of Frame m ( $m > 1$  and m are the natural number) -- seeing -- " -- visual information is continued and inputted like and the hit of Frame n ( $n > m$  and m are the natural number) is seen further, and when touching "left ankle, it is made to touch -- left ankle If visual information is continued and inputted so that it may be ", the visual information over an animation can be given subjectively.

[0066] Then, retrieval actuation of an image is explained. The flow chart with which drawing 12 explains retrieval actuation of the gestalt 1 of operation, and drawing 13 are drawings showing an example of the display transition at the time of retrieval actuation of the gestalt 1 of operation. Here, the example which searches the image a mentioned above from an image database 5 is given and explained.

[0067] If processing shifts to retrieval processing, as shown in drawing 13 (a), the image display area 301 and the natural language display area 302 will be displayed on the screen of a display 4. And as shown in drawing 13 (a), the message (please input "retrieval sentence in the example of drawing") 303 of a purport which inputs the retrieval sentence about the image used as the object for retrieval with natural language is displayed on the natural language display area 302 (step S61).

[0068] In addition, the activation icon 305 for directing the termination icon 304 for carrying out termination actuation of this retrieval processing and activation of retrieval is displayed in the natural language display area 302.

[0069] Then, a retrieval sentence, i.e., a natural language entry of data, is received (step S62). Natural language data are data inputted through the natural language input section 2. Initiation of a natural language input displays a retrieval sentence in the natural language display area 302 like "the image which has placed the drink which got cold on the table is ?", as shown in drawing 10 (b).

[0070] Also in this case, although not illustrated, it is displayed as Cursor CUR shows the input location and correction location of the following, and migration actuation of that cursor CUR is carried out by the control unit which is not illustrated. In addition, the activation directions (step S63) by the activation icon 305 and the termination directions (step S64) by the termination icon 304 are sensed until the input of a retrieval sentence is completed.

[0071] And in the language analysis section 7, if the input of a retrieval sentence is completed and selection actuation (the slash in drawing shows) of the activation icon 305 is carried out as shown in drawing 13 (b) (step S63), while language analysis based on the natural language data of an input retrieval sentence is performed (step S65), language analysis based on the natural language data of the visual information on an image database 5 will be performed (step S66).

[0072] Thus, collating is performed where the base is arranged by carrying out language analysis of both input retrieval sentence and visual information (step S67). This collating follows the principle explained by above-mentioned drawing 3. If it furthermore explains in full detail, the impression common to the information as which this collating is expressed in a retrieval sentence on the context of natural language and in visual information will judge which is contained. Here, as long as it is a collating technique by similarity (whenever [ coincidence ]), what kind of existing technique shall be applied.

[0073] Thus, if a collating result is obtained, the high file of similarity will be extracted. in this case, file name "\*\*\*\* which holds Image a -- the file of that 1" is extracted. Therefore, as shown in drawing 13 (c), based on the image data of the file, Image a is displayed on the image display area 301 (step S68).

[0074] here, since the high image of similarity is displayed, if you may make it display a number of images beforehand decided to be ascending order when there were two or more candidates, and two or more images are displayed as a candidate and a desired image is in it, display only the image of the request -- it may obtain and you may make it like. Since these techniques are well-known, the explanation is omitted here.

[0075] What is necessary is just to operate the activation icon 305 after step S68, after changing a retrieval sentence if the image of further others is searched since processing returns to step S62 (step S62 and step S63). Moreover, retrieval processing can be terminated by selection actuation of the termination icon 304 (step S64).

[0076] Here, an example at the time of realizing an above-mentioned image processing system by hardware is given. Drawing 14 is the block diagram showing the image processing system by the gestalt 1 of implementation of this invention in hardware.

[0077] In the hardware configuration shown in drawing 14, also functionally, a display 4 and an image database 5 share also in hardware, and also a scanner 111, a keyboard 121, a mouse 122, CPU131, program memory 132, and RAM133 are applied. It connects with Bus BS and each [ these ] unit receives control of CPU131 through Bus BS.

[0078] CPU131 is a unit which operates according to the various programs of program memory 132, and performs processing which is equivalent to an image / natural language correlation section 3, the retrieval section 6, and the language analysis section 7 in functional block of drawing 4 R> 4. Program memory 132 stores operating system OS132A and application program 132B for CPU131 to operate. RAM133 is used as a work area of CPU131.

[0079] In addition, the dictionary section 8 in functional block of drawing 4 may be connected to Bus BS as dictionary memory which could be stored in a part of application program 132B, or became independent.

[0080] A scanner 111 is a unit equivalent to the image input section 1, and outputs the image data inputted from the manuscript image to RAM133 by control of CPU131. A keyboard 121 is a unit equivalent to the natural language input section 2, inputs an alphabetic character etc. and inputs data, such as a visual information and retrieval sentence. A mouse 122 is a pointing device which supports actuation of the input of a keyboard 121, each above-mentioned icon selection, etc.

[0081] Since according to the gestalt 1 of operation the natural language data in which the visual impressions of the image formed based on image data and its image data are shown were associated and it memorized to the image database

5 as explained above, it is possible to put in a database with the subjective element in which visual impressions have an image.

[0082] Moreover, the natural language data in which the visual impressions of the image formed in an image database 5 based on image data and its image data are shown are associated and memorized. Since the image data relevant to the natural language data was made to answer when access was performed by the natural language data of a certain visual impressions on an image database 5 Regardless of attribute information, such as classification of filing time and an image, access by visual impressions, i.e., a subjective impression, is attained, and it becomes easy to access a desired image by this.

[0083] Moreover, since image data was used as data, such as still picture data and a video data, it is possible to give the visual information on arbitration to each image regardless of the classification of an image.

[0084] Moreover, the natural language data in which the visual impressions about the image formed in an image database 5 based on image data and its image data are shown are associated and memorized. By inputting the natural language data in which a retrieval sentence is shown, and collating with the natural language data on an image database 5 Since the image data relevant to the natural language data with which visual impressions are common was extracted from the image database 5 It is possible for an image to be searched with visual impressions, i.e., a subjective impression, even if it does not know the attribute information on an image, and to search a desired image more certainly and efficiently by this.

[0085] Moreover, since it was made to collate after carrying out language analysis of the natural language data inputted as the natural language data memorized by the image database 5 on the occasion of retrieval, high collating of the precision suitable for the semantics of visual impressions is possible.

[0086] Moreover, since it was made to perform image display based on the image data extracted by retrieval, it is possible for a retrieval result to be shown visually and to check the correction of a retrieval result easily by this.

[0087] Moreover, it is possible to give the visual information on arbitration to each image, without being caught by the objectivity of an image, since the natural language data for relating with image data and memorizing were inputted through the key stroke.

[0088] (Gestalt 2 of operation) Now, after carrying out language analysis of the natural language data which are a retrieval sentence, and the natural language data on an image database 5, respectively, he was trying to collate with the gestalt 1 of operation mentioned above, but like the gestalt 2 of the operation explained below, after reasoning visual impressions about the natural language data on an image database 5, it may be made to collate.

[0089] In this case, since the whole configuration (refer to drawing 4 ) by the gestalt 1 of the above-mentioned operation is used for the gestalt 2 of operation about a whole configuration, it omits explanation of activity \*\*\*\* for a common number about an intersection. The configuration and function which are different from below are explained, and it omits about a common feature.

[0090] First, a configuration is explained. Drawing 15 is the block diagram showing functionally the image processing system by the gestalt 2 of implementation of this invention. Also with the gestalt 2 of operation, a principle becomes being the same as that of the gestalt 1 of the above-mentioned operation. With the gestalt 2 of this operation, the inference section 9 and the information database 10 are newly added to an image processing system. Although inference is actually carried out also in the language analysis section 7, it uses the inference section 9 and the information database 10 as a component here in respect of generation of the new visual information between the visual information on the plurality after language analysis.

[0091] The inference section 9 makes the new visual information which can predict the analysis result of the language analysis section 7 from two or more visual information to the analysis result of a retrieval sentence with reference to reception and the information database 10, and outputs it to the retrieval section 6 by making this into an inference result. The information database 10 offers exchange for combining the analysis result of reception and two or more visual information for an analysis result from all views, and making a new subjective impression from the inference section 9.

[0092] In addition, on a hardware configuration, the inference section 9 shall be equivalent to CPU131, and the information database 10 may be the memory configuration which became independent like the dictionary section 8 at a part of application program AP132B.

[0093] Below, the retrieval actuation which is different from the gestalt 1 of operation is explained. Drawing 16 is a flow chart explaining the important section of retrieval actuation of the gestalt 2 of operation, and drawing 17 is drawing showing an example of the display transition at the time of retrieval actuation of the gestalt 2 of operation. In drawing 16, the graphic display of the box which is common in drawing 12 is omitted, and is illustrated only about the box added to it.

[0094] Also in the gestalt 2 of this operation, like the gestalt 1 of operation mentioned above, if retrieval processing is

started, the image display area 401 and the natural language display area 402 will be displayed on a display 4 (refer to drawing 17 (a)). Since message "inputting retrieval sentence" 403 are displayed in the natural-language display area 402 at this time, if selection actuation of the activation icon 405 is carried out after the input of the retrieval sentence "the image of the liquid which will get drunk if it drinks is ?", according to a retrieval sentence "the image of the liquid which will get drunk if it drinks is ?", a search will be performed in the natural-language display area 402 (refer to drawing 17 (b)).

[0095] After the language analysis of the visual information on an image database 5 is completed in step S66 [ after the language analysis of the above-mentioned retrieval sentence is completed in step S65 ], reasoning is performed with the gestalt 2 of this operation. That is, based on the analysis result of visual information, the content near the analysis result of an input retrieval sentence is reasoned (step S69). This principle is explained briefly below.

[0096] Here, the above-mentioned images a, b, and c are mentioned as an example. In the case of Image a, as shown in drawing 1, when it was Image c, as the analysis result of the image of "wine" was obtained, it was shown in drawing 3 R> 3, and the analysis result of the image of "Biel" was obtained, and it was shown in drawing 3, the analysis result of "coffee" is obtained [ when it was Image b, ].

[0097] Then, if "wine contains alcohol", "Biel containing alcohol", "coffee not containing alcohol", and "drunkenness will turn if the drink containing alcohol is drunk" are registered into the information database 10, inference that it is the liquid which will get drunk if "wine" is drunk about Image a is materialized. Inference that it is the liquid which similarly will get drunk if "Biel" is drunk about Image b is materialized. On the other hand, inference of the liquid which does not get drunk although "coffee" is drunk about Image c is materialized.

[0098] Consequently, the visual information on a content like "seeming to be the image of the wine which gets drunk, if it drinks" is newly generated about Image a, the visual information on a content like "seeming to be the image of Biel which gets drunk, if it drinks" is newly generated about Image b, and the visual information on a content like "seeming to be the image of the coffee which does not get drunk, although drunk" is newly generated about Image c. These new visual information is outputted to the retrieval section 6 from the inference section 9 in the condition of the analysis result stated with the gestalt 1 of operation for collating of step S67.

[0099] Henceforth, image display (step S68) by collating (step S67) in the condition of the analysis result of an input retrieval sentence and visual information and the extract of the high file of similarity is performed like the gestalt 1 of operation mentioned above. Consequently, as shown in drawing 17 (c), Image a and Image b are displayed as an object for retrieval, and become the image display area 401 the outside of an object about Image c.

[0100] As explained above, according to the gestalt 2 of operation, the natural language data in which two or more visual impressions about the image formed in an image database 5 based on image data and its image data are shown are associated and memorized. When the new visual impressions equivalent to the visual impressions of the natural language data which are a retrieval sentence are able to be drawn from two or more visual impressions of a certain natural language data on an image database 5, referring to the information database 10 Since the image data relevant to the natural language data of a certain was extracted from the image database 5 It is possible to be able to secure as an object for an extract of an image certainly, even if it is natural language data which are overlooked, when the visual information on an image database 5 is used independently, and to search a desired image more certainly and efficiently by this in the case of retrieval.

[0101] Moreover, since it was made to reason after carrying out language analysis of the natural language data inputted as the natural language data memorized by the image database 5 on the occasion of inference, high inference of the precision suitable for the semantics of visual impressions is possible.

[0102] Moreover, since ascending order extracted two or more image data from the higher one of similarity according to the collating result on the occasion of retrieval, it is possible for two or more images near visual impressions to be mentioned as a candidate, and to search a desired image more efficiently even from visual impressions by this.

[0103] (Gestalt 3 of operation) Now, although he was trying to input natural language data by manual actuation through the natural language input section 2, you may make it automate a natural language entry of data like the gestalt 3 of the operation explained below with the gestalten 1 and 2 of operation mentioned above using an image (graphic form) recognition technique.

[0104] In this case, since the whole configuration (refer to drawing 4 or drawing 15 ) by the gestalten 1 or 2 of the above-mentioned operation is used for the gestalt 3 of operation about a whole configuration, it omits explanation of activity \*\*\*\* for a common number about an intersection. Only the configuration and function which are different from below are explained, and it omits about a common feature.

[0105] First, the configuration which is different from the gestalten 1 and 2 of operation is explained. It is the block diagram showing the important section of the image processing system by the gestalt 3 of implementation of invention of

**\*\*18\*\*** in hardware. With the gestalt 3 of operation, the image recognition program memory 125 which newly stored the image recognition program, and the image recognition dictionary 126 which supports image recognition processing are connected to Bus BS. Detailed explanation is omitted for the well-known image recognition technique by these addition configuration.

[0106] Below, the filing actuation which is different from the gestalten 1 and 2 of operation is explained. Drawing 19 is a flow chart explaining filing actuation of the gestalt 3 of operation. Termination of an image input (step S51) and input image display (step S52) performs recognition of an image based on the image data inputted with reference to the image recognition dictionary 126 like the gestalten 1 and 2 of the above-mentioned operation according to an image recognition program (step S57). It becomes an element at the time of physical relationship with the configuration of each image component, size, a color, and other image components etc. recognizing in that case.

[0107] Furthermore, the visual information on an image is extracted from the recognition result of an image (step S58). What is necessary is to extract use of color and the visual information of [ if a color scheme is dark ] "the image of dark sensibility" for example, and just to, extract the visual information of "the image of bright sensibility" on the other hand in this extract, for example, if use of color and a color scheme are bright in order to acquire a subjective impression. If visual information is extracted by such regularity, since it will be avoided that the impression of an image is restricted by objectivity, a more subjective impression can be related with an image like a key input or voice input.

[0108] Pictures are one of those to which this subjective impression is applied. It is possible to give a forcible image, a delicate image, and various expressions from the touch of the brush of that man proper being expressed by the picture as it is in the case of these pictures. For this reason, the application to the utilization gestalt of managing a book of paintings is possible.

[0109] Moreover, when it glances in an image and two persons and a mirror exist, you may make it give the content "there are two persons" or "one person is a actual person and one more person is a person in a mirror" as visual information. Then, if the visual information which contains the latter "mirror" to the above-mentioned image is matched, it is possible to search for example, also about the image in which two persons exist according to inference that people turn into two persons, in a "mirror" even if the retrieval conditions "people are one-person \*\*\*\* surely" are given.

[0110] Henceforth, a response with the visual information and the image data which were obtained at step S58 as well as the gestalten 1 and 2 of operation mentioned above is associated, and registration to an image database 5 is performed (step S56).

[0111] It is possible to give the visual information which has subjectivity with each image, without being caught by the objectivity of an image, if the image recognition dictionary 126 is constituted so that the visual information which expresses subjectively the color of an image, the configuration of each body in an image, size, etc. may be acquired since the natural language data for relating with image data and memorizing were inputted through image recognition according to the gestalt 3 of operation as explained above.

[0112] (Gestalt 4 of operation) Now, although one image processing system is made to perform filing and retrieval, an image processing system is divided into a client and a server, and you may make it distribute a function like the gestalt 4 of the operation explained below with the gestalten 1-3 of operation mentioned above.

[0113] Here, only a configuration is explained. Drawing 20 is a block diagram showing functionally the image processing system by the gestalt 4 of implementation of this invention. The image processing system by the gestalt 4 of operation is divided into a client 21 and a server 22 as shown in drawing 20. Although image data and natural language data related with the image database 5 and it memorized with the gestalt of each above-mentioned operation, with the gestalt 4 of this operation, the natural language data which are visual information are memorized by the client 21, and the image data related with the server 22 by that visual information is memorized.

[0114] These visual information and image data are associated by adding the search keys (address, identification information, etc.) for accessing image data to visual information. That is, a search key common to visual information and image data is added.

[0115] Therefore, in the case of retrieval, if collating with the input retrieval sentence after language analysis and visual information is performed like the gestalt of each operation first mentioned above and it becomes the phase of an image extract after that, a search key will be taken out from the high visual information on similarity, and desired image data will be extracted from a server 22 by the search key. A search key and the image data which it is as a result of retrieval are delivered and received through a transmission line in between a client 21 and a server 22.

[0116] In addition, in the case of filing, after a client 21 goes to correlation containing a search key, it leaves visual information to self-equipment and transmits to a server 22 about image data. A server 22 memorizes the image data which received from the client 21.

[0117] Since it becomes unnecessary to memorize [ according to the gestalt 4 of operation ] image data by the client side

by distributing a function in addition to each effectiveness of the gestalten 1-3 of operation as explained above, it is possible to hold down the amount of data of a client side to necessary minimum.

[0118] (Gestalt 5 of operation) Now, with the gestalt 1 of operation mentioned above, although the natural language input by the keyboard 121 was mentioned as the example, voice input may be made to perform a natural language input like the gestalt 5 of the operation explained below. Here, since it is only the difference in the input configuration of natural language, only the point is explained below.

[0119] Drawing 21 is the block diagram showing the important section of the image processing system by the gestalt 5 of implementation of this invention in hardware. In drawing 21, the speech recognition program memory 124 which stored the speech recognition program for performing speech recognition based on the microphone 123 which inputs voice, and the voice data inputted from the microphone 123 is connected to Bus BS.

[0120] This microphone 123 and the speech recognition program memory 124 are added to the configuration of the gestalt of each above-mentioned operation as other natural language input sections from constituting the natural language input section 2. Natural language data can be inputted from voice by this microphone 123 and the speech recognition program memory 124.

[0121] Since the natural language data for relating with image data and memorizing were inputted through speech recognition according to the gestalt 5 of operation as explained above, it is possible like the gestalt 1 of the above-mentioned operation to give the visual information on arbitration to each image, without being caught by the objectivity of an image.

[0122] (Gestalt 6 of operation) Now, with the gestalt 1 of operation mentioned above, although the image input with a scanner 111 was mentioned as the example, a communication link may be made to perform an image input like the gestalt 6 of the operation explained below. Here, since it is only the difference in an image entry-of-data configuration, only the point is explained below.

[0123] Drawing 22 is the block diagram showing the important section of the image processing system by the gestalt 6 of implementation of this invention in hardware. In drawing 22, communication link I/F112 connected to a communication line and the modem 113 which carries out the strange recovery of the data transmitted and received through the communication link I/F112 are connected to Bus BS.

[0124] This communication link I/F112 and modem 113 are added to the configuration of the gestalt of each above-mentioned operation as other image input sections from constituting the image input section 1. This communication link I/F112 and modem 113 can be boiled, and can input image data from the exterior more.

[0125] If visual information is given to a homepage in recent years on the occasion of utilization of the Internet which is spreading also for the application which captures an image through a communication line according to the gestalt 6 of operation since the same effectiveness as the gestalten 1-5 of the above-mentioned operation is acquired as explained above, it will become possible to manage a huge homepage easily only by the visual impression.

[0126]

[Effect of the Invention] Since the natural language data in which the visual impressions of the image formed based on image data and its image data are shown were associated and memorized according to invention of claim 1 as explained above, the effectiveness that the image management equipment which can be put in a database with the subjective element in which visual impressions have an image is obtained is done so.

[0127] Moreover, according to invention of claim 2, the natural language data in which the visual impressions of the image formed in an image database based on image data and its image data are shown are associated and memorized. Since the image data relevant to the natural language data was made to answer when access was performed by the natural language data of a certain visual impressions on an image database The effectiveness that the image management equipment with which access by visual impressions, i.e., a subjective impression, is attained regardless of attribute information, and it becomes easy to access a desired image by this is obtained is done so.

[0128] Moreover, the effectiveness that the image management equipment which can give the visual information which has subjectivity in each image is obtained is done so, without being caught by the objectivity of an image, since the natural language data for relating with image data and memorizing in invention of claims 1 or 2 were inputted through a key stroke, speech recognition, image recognition, etc. according to invention of claim 3.

[0129] Moreover, according to invention of claim 4, in any one invention of claims 1-3, since image data was used as data, such as still picture data and a video data, the effectiveness that the image management equipment which can give the visual information on arbitration to each image is obtained regardless of the classification of an image is done so.

[0130] Moreover, according to invention of claim 5, the natural language data in which the visual impressions about the image formed in an image database based on image data and its image data are shown are associated and memorized. By inputting the natural language data in which a retrieval sentence is shown, and collating with the natural language data on



an image database Since the image data relevant to the natural language data with which visual impressions are common was extracted from the image database The effectiveness that the image retrieval equipment which an image is searched with visual impressions, i.e., a subjective impression, even if it does not know the attribute information on an image, and can search a desired image more certainly and efficiently by this is obtained is done so.

[0131] Moreover, since it was made to collate according to invention of claim 6 after carrying out language analysis of the natural language data inputted as the natural language data memorized by the image database on the occasion of retrieval in invention of claim 5, the effectiveness that the image retrieval equipment in which high collating of the precision suitable for the semantics of visual impressions is possible is obtained is done so.

[0132] Moreover, according to invention of claim 7, in invention of claims 5 or 6, since ascending order extracted two or more image data from the higher one of similarity according to the collating result on the occasion of retrieval, two or more images near visual impressions are mentioned as a candidate, and the effectiveness that the image retrieval equipment which can search a desired image more efficiently even from visual impressions by this is obtained is done so.

[0133] Moreover, according to invention of claim 8, the natural language data in which two or more visual impressions about the image formed in an image database based on image data and its image data are shown are associated and memorized. When the new visual impressions equivalent to the visual impressions of the natural language data which are a retrieval sentence are able to be drawn from two or more visual impressions of a certain natural language data on an image database, referring to an information database Since the image data relevant to the natural language data of a certain was extracted from the image database When the visual information on an image database is independently used on the occasion of retrieval, even if it is natural language data which are overlooked, are certainly securable as an object for an extract of an image. By this The effectiveness that the image retrieval equipment which can search a desired image more certainly and efficiently is obtained is done so.

[0134] Moreover, since it was made to reason according to invention of claim 9 after carrying out language analysis of the natural language data inputted as the natural language data memorized by the image database on the occasion of inference in invention of claim 8, the effectiveness that the image retrieval equipment in which high inference of the precision suitable for the semantics of visual impressions is possible is obtained is done so.

[0135] Moreover, since it was made to perform image display in any one invention of claims 5-9 based on the image data extracted by retrieval according to invention of claim 10, a retrieval result is shown visually and the effectiveness that the image retrieval equipment which can check the correction of a retrieval result easily by this is obtained is done so.

[0136] Moreover, since according to invention of claim 11 it was made the process which memorizes image data and natural language data after associating the natural language data in which the visual impressions of the image formed based on image data and this image data are shown, the effectiveness that the image management method which can be put in a database with the subjective element in which visual impressions have an image is acquired is done so.

[0137] Moreover, according to invention of claim 12, the natural language data in which a retrieval sentence is shown are inputted. The natural language data inputted to the natural language data memorized by the image database which associated and memorized the natural language data in which the visual impressions about the image formed based on image data and its image data are shown are collated. Since the image data relevant to the natural language data with which visual impressions are common according to the collating result was made into the process extracted from an image database The effectiveness that the image search method which an image is searched with visual impressions, i.e., an objective impression, even if it does not know the attribute information on an image, and can search a desired image more certainly and efficiently by this is obtained is done so.

[0138] Moreover, according to invention of claim 13, by having recorded the program which makes a computer perform the approach indicated by claims 11 or 12, reading in a computer becomes possible about the program, and the effectiveness that the record medium which can realize actuation of claims 11 or 12 by computer is obtained is done so by this.

---

[Translation done.]



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**